Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

Mon. Not. R. Astron. Soc. **406,** 612–625 (2010) doi:10.1111/j.1365-2966.2010.16701.x

A first direct measurement of the intergalactic medium temperature around a quasar at $z = 6$

James S. Bolton,[1⋆] George D. Becker,[2] J. Stuart B. Wyithe,[1] Martin G. Haehnelt[2] and Wallace L. W. Sargent[3]

[1]*School of Physics, University of Melbourne, Parkville, VIC 3010, Australia*
[2]*Kavli Institute for Cosmology and Institute of Astronomy, Madingley Road, Cambridge CB3 0HA*
[3]*Palomar Observatory, California Institute of Technology, Pasadena, CA 91125, USA*

Accepted 2010 March 17. Received 2010 March 17; in original form 2010 January 14

ABSTRACT

The thermal state of the intergalactic medium (IGM) provides an indirect probe of both the H I and He II reionization epochs. Current constraints on the IGM temperature from the Lyα forest are restricted to the redshift range $2 \leq z \leq 4.5$, limiting the ability to probe the thermal memory of H I reionization towards higher redshift. In this work, we present the first direct measurement of the IGM temperature around a $z = 6$ quasar by analyzing the Doppler widths of Lyα absorption lines in the proximity zone of SDSS J0818+1722. We use a high-resolution ($R = 40\,000$) Keck/HIRES spectrum in combination with detailed numerical modelling to obtain the temperature at mean density, $T_0 = 23\,600 \pm {}^{5000}_{6900}$ K ($\pm^{9200}_{9300}$ K) at 68 (95) per cent confidence assuming a prior probability $13\,500\,\mathrm{K} \leq T_0 \leq 38\,500\,\mathrm{K}$ following H I and He II reionization. This enables us to place an upper limit on the redshift of H I reionization, z_H, within 33 comoving Mpc of SDSS J0818+1722. If the quasar reionizes the He II in its vicinity, then in the limit of instantaneous reionization we infer $z_\mathrm{H} < 9.0(11.0)$ at 68 (95) per cent confidence assuming photoheating is the dominant heat source and that H I reionization is driven by ionizing sources with soft spectra, typical of Population II stars. If the H I and He II in the IGM around SDSS J0818+1722 are instead reionized simultaneously by a population of massive metal-free stars, characterized by very hard ionizing spectra, we obtain a tighter upper limit of $z_\mathrm{H} < 8.4(9.4)$. Initiating reionization at higher redshifts produces temperatures which are too low with respect to our constraint unless the H I ionizing sources or the quasar itself has spectra significantly harder than typically assumed.

Key words: methods: numerical – intergalactic medium – quasars: absorption lines.

1 INTRODUCTION

The discovery of $z \simeq 6$ quasars within the last decade has led to several important advances in our understanding of the high-redshift Universe. One area where the impact of this work has been especially significant is the study of the high-redshift intergalactic medium (IGM) with quasar absorption lines. The Lyα forest in particular provides a valuable probe of the ionization state of the IGM (Fan et al. 2002, 2006; Songaila 2004; Becker, Rauch & Sargent 2007). However, the increasing opacity of the IGM to Lyα photons, culminating in the appearance of the Gunn & Peterson (1965) trough at $z \simeq 6$ (Becker et al. 2001; White et al. 2003), ultimately limits the utility of the Lyα forest at the highest observable redshifts. This limitation has led many authors to analyze the small regions which exhibit transmission through the Lyα forest even at $z > 6$. These highly ionized proximity zones lie between the quasar Lyα emission line and the red-most edge of the Gunn–Peterson trough, and are due to the enhanced ionization of hydrogen close to the quasar.

Previous analyses have focused on using these regions to examine the ionization state of the IGM with low- to moderate-resolution spectra ($R \sim 3000$–6000). These studies have typically modelled the extent and/or shape of the observed transmission to obtain constraints on the IGM ionization state at $z \simeq 6$ (e.g. Fan et al. 2006; Mesinger & Haiman 2007; Alvarez & Abel 2007; Bolton & Haehnelt 2007a; Wyithe, Bolton & Haehnelt 2008; Maselli, Ferrara & Gallerani 2009). However, the proximity zones can also be used to probe the *thermal state* of the IGM at high redshift. High-resolution ($R \sim 40\,000$) spectra resolve the thermal broadening kernel, enabling the widths of Lyα absorption lines, and hence the temperature of the gas in the proximity zones, to be directly

⋆E-mail: jsbolton@unimelb.edu.au

measured (Becker, Sargent & Rauch 2005). The linewidths will be sensitive to any photoheating caused by the quasar itself (Miralda-Escudé & Rees 1994; Bolton & Haehnelt 2007b). Furthermore, since the cooling time-scale in the low-density IGM is long, information on the reionization history *prior* to any quasar activity will also be encoded in these absorption lines (Haehnelt & Steinmetz 1998; Theuns et al. 2002a; Hui & Haiman 2003).

Earlier measurements of the IGM thermal state using the Lyα forest at $2 \leq z \leq 4.5$ have indeed yielded valuable insights into the epoch of He II reionization at $z \simeq 3$ (Schaye et al. 2000; Ricotti, Gnedin & Shull 2000; McDonald et al. 2001; Zaldarriaga 2002; Lidz et al. 2009). However, given enough time following reionization, the IGM temperature will eventually reach an asymptotic value which depends only on the spectral shape of the ionizing background. Information on the earlier IGM thermal evolution, and hence the H I reionization history at high redshift, will then be effectively erased (Theuns et al. 2002a; Hui & Haiman 2003). Consequently, it is desirable to study the thermal state of the IGM as close as possible to the H I reionization epoch, where the temperature retains a more recent memory of H I photoheating (Cen et al. 2009; Furlanetto & Oh 2009). The linewidths in proximity zones at $z \simeq 6$ thus contain valuable information on the thermal history of the IGM at $z \gtrsim 6$ as well as the impact of any photoheating by the quasar itself on its environment (Bolton & Haehnelt 2007b; Lidz et al. 2007).

In this paper, we perform an analysis of Lyα absorption linewidths in the proximity zone of the $z = 6$ quasar SDSS J0818+1722. We compare high-resolution Keck/HIRES data with detailed synthetic spectra constructed using hydrodynamical simulations and line-of-sight radiative transfer (RT) to obtain the first direct constraint on the IGM temperature around a quasar at $z \simeq 6$. As such, this work represents a first step towards developing a procedure which may be applied to larger data sets at high redshift in the future. We begin in Section 2 by introducing the observational data and numerical simulations we use for our analysis. We discuss our methodology in Section 3, and examine the systematic uncertainties which may impact on our results in Section 4. We present our temperature measurement in Section 5. In Section 6, we consider the implications of our results for the IGM reionization history around SDSS J0818+1722 before finally concluding in Section 7. All distances are expressed in comoving units unless otherwise stated.

2 DATA AND NUMERICAL MODELLING

2.1 Observational data

Our temperature measurements are based on a high-resolution spectrum of the $z = 6.00$ quasar SDSS J0818+1722 (Fan et al. 2006). Keck HIRES data were taken in 2006 February using the upgraded detector. We employed a 0.86 arcsec slit, which produces a resolution of $R = 40\,000$ [full width at half-maximum (FWHM) = 6.7 km s^{-1}]. The total integration time was 7.5 h. The raw data were processed using a custom set of IDL routines that include optimal sky subtraction (Kelson 2003). In order to achieve the highest possible signal-to-noise ratio (S/N) ratio, the final one-dimensional spectrum was optimally extracted from all exposures simultaneously after relative flux calibrations were applied to remove the blaze functions from individual orders. This allowed us to efficiently reject cosmic rays and other bad pixels while preserving as many counts from the object as possible. The final S/N near the Lyα emission line was $\sim$15 per 2.1 km s^{-1} binned pixel.

Continuum fitting in the proximity zone region of a $z \sim 6$ quasar is challenging due to the strong absorption present in the blue side of the Lyα emission line. We first divided the combined spectrum by a power law with $F_\nu \propto \nu^{-0.5}$, normalized near $1280(1 + z)$ Å. The value of the power-law slope only weakly affects the overall continuum level in the proximity zone. Next, we fit a slowly varying spline to the emission line. The blue side of the line profile was made to roughly mirror the red side, which could be drawn through the unabsorbed continuum. This produced a generally Gaussian shape, which we joined smoothly on to the underlying power law. The continuum was set $\sim$10 per cent above the top of the transmission peaks within $\sim$2000 km s^{-1} of the quasar redshift, after which the peaks were allowed to fall further below the continuum. Although the continuum level is obscured by the high levels of absorption, we estimate that our fit is within $\sim$20 per cent of the correct value over the wavelength range of interest.

2.2 Hydrodynamical simulations

The synthetic proximity zone spectra used in this study are constructed using line-of-sight density, velocity and temperature fields drawn from cosmological hydrodynamical simulations. The simulations were performed using a customized version of the parallel Tree-SPH code GADGET-3, which is an updated version of the publicly available code GADGET-2 (Springel 2005). We use 12 different hydrodynamical simulations in this work, summarized in Table 1. Nine simulations (A–H and W) were performed in $10\,h^{-1}$ Mpc periodic boxes containing 2×512^3 gas and dark matter particles. The simulations are specifically designed to resolve the Lyα forest at high redshift (Bolton & Becker 2009). However, they employ a relatively small box size to achieve the required mass resolution. In order to assess the effect of box size on our results, we performed two further simulations with identical mass resolution but different box sizes of $10\,h^{-1}$ Mpc (L1) and $40\,h^{-1}$ Mpc (L3), respectively. Finally, another $10\,h^{-1}$ Mpc simulation, L2, was performed with 2×256^3 gas and dark matter particles to provide an additional test of convergence with mass resolution.

The simulations were all started at $z = 99$, with initial conditions generated using the transfer function of Eisenstein & Hu (1999). The cosmological parameters are $\Omega_\mathrm{m} = 0.26$, $\Omega_\Lambda = 0.74$, $\Omega_\mathrm{b} h^2 = 0.023$, $h = 0.72$, $\sigma_8 = 0.80$, $n_\mathrm{s} = 0.96$, consistent with recent studies of the cosmic microwave background (Komatsu et al. 2009; Reichardt et al. 2009). The IGM is assumed to be of primordial composition with a helium fraction by mass of $Y = 0.24$ (Olive & Skillman 2004). The gravitational softening length was set to 1/30th of the mean linear interparticle spacing and – with the exception of model W – star formation was included using a simplified prescription which converts all gas particles with overdensity $\Delta = \rho/\langle\rho\rangle > 10^3$ and temperature $T < 10^5$ K into collisionless stars. Model W instead uses the multi-phase star formation and feedback model of Springel & Hernquist (2003). We use this to explore the impact of strong galactic winds on our results.

The gas in the simulations is assumed to be optically thin and in ionization equilibrium with a spatially uniform ultraviolet background (UVB). The UVB corresponds to the galaxies and quasars emission model of Haardt & Madau (2001). Hydrogen is reionized at $z = 9$ and gas with $\Delta \lesssim 10$ subsequently follows a tight power-law temperature–density relation, $T = T_0 \Delta^{\gamma-1}$, where T_0 is the temperature of the IGM at mean density (Hui & Gnedin 1997; Valageas, Schaeffer & Silk 2002). In order to explore a variety of thermal histories, we rescale the Haardt & Madau (2001) photoheating rates by different constants in models A–H. In each simulation

Table 1. Hydrodynamical/RT simulations used in this work. From left to right, the columns list the simulation identifier, the box size, the total particle number, the gas particle mass, the scaling factors for the UVB photoheating rates (see main text for details), the median volume-weighted gas temperature at mean density, T_0, and the power-law slope of the temperature–density relation, $T = T_0\Delta^{\gamma-1}$, both prior to and after He II photoheating by the quasar. The slope is given in terms of $\gamma - 1 = 2(\log T_0 - \log T_{-0.5})$, where $T_{-0.5}$ is the median volume-weighted gas temperature at $\log\Delta = -0.5$. The temperatures are quoted to three significant figures only. Note a power-law temperature–density relation following He II photoheating by the quasar is an approximation only (see Fig. 1).

Model	$L(h^{-1}$ Mpc)	Particles	$M_{\rm gas}(h^{-1}\,{\rm M}_\odot)$	ζ	ξ	$T_{0,\rm i}$ (K)	$\gamma_{\rm i} - 1$	T_0 (K)	$\gamma - 1$
A	10	2×512^3	9.2×10^4	0.30	0.0	4 300	0.39	13 500	0.12
B	10	2×512^3	9.2×10^4	0.80	0.0	8 500	0.40	17 500	0.18
C	10	2×512^3	9.2×10^4	1.45	0.0	13 000	0.41	21 500	0.22
D	10	2×512^3	9.2×10^4	2.20	0.0	17 300	0.41	25 700	0.25
E	10	2×512^3	9.2×10^4	3.10	0.0	21 900	0.41	29 900	0.27
F	10	2×512^3	9.2×10^4	4.20	0.0	26 800	0.41	34 400	0.28
G	10	2×512^3	9.2×10^4	5.30	0.0	31 100	0.40	38 500	0.29
H	10	2×512^3	9.2×10^4	1.45	-1.0	12 800	0.06	21 400	0.04
L1	10	2×128^3	5.9×10^6	1.45	0.0	13 600	0.40	22 300	0.22
L2	10	2×256^3	7.3×10^5	1.45	0.0	13 200	0.40	22 000	0.21
L3	40	2×512^3	5.9×10^6	1.45	0.0	13 400	0.41	20 500	0.23
W	10	2×512^3	9.2×10^4	1.45	0.0	13 000	0.41	21 700	0.22

we assume $\epsilon_i = \zeta\,\Delta^{\xi}\epsilon_i^{\rm HM01}$, where $\epsilon_{\rm i}^{\rm HM01}$ are the Haardt & Madau (2001) photoheating rates for species i = [H I, He I, He II] and ζ, ξ are constants listed in Table 1. The L1, L2, L3 and W simulations use same thermal history as model C. A density-dependent rescaling of the photoheating rates has been applied to model H, giving $\gamma \simeq 1$ while maintaining a similar T_0 to model C. All the other hydrodynamical simulations in this study have $\gamma \simeq 1.4$. Simulation outputs were obtained from each model at $z = 6.01$. The values of T_0 and γ in the hydrodynamical simulations are listed in Columns 7 and 8 of Table 1.

2.3 Line-of-sight radiative transfer models

In order to correctly model photoionization and heating[1] by the quasar, we have implemented a line-of-sight RT scheme in the hydrodynamical simulations to account for the propagation of ionizing radiation into the IGM. For this we use the multi-frequency line-of-sight RT algorithm described by Bolton & Haehnelt (2007b) and updated in Bolton, Oh & Furlanetto (2009).

We follow the procedure outlined in Bolton & Haehnelt (2007b) for constructing the initial conditions for our simulations including RT. We use a friends-of-friends halo finding algorithm with a linking length of 0.2 to identify haloes in the hydrodynamical simulations. We then select the 10 most massive haloes in each simulation and extract lines-of-sight in different orientations around them. The exact halo mass has little impact; we find it is more important to resolve the IGM along the line-of-sight rather than the host halo itself. We discuss this issue further in Section 4.1. The halo lines-of-sight are then spliced with lines-of-sight drawn at random from the rest of the simulation, resulting in 30 different lines-of-sight $55\,h^{-1}$ Mpc in length, all of which start at the location of a halo.

For each line-of-sight we compute the transfer of ionizing radiation through the IGM from a quasar which emits $\dot{N}$ photons per second above the H I ionization threshold,

$$\dot{N} = \int_{\nu_{\rm H\,I}}^{\infty} \frac{L_\nu}{h_{\rm p}\nu}\,{\rm d}\nu, \tag{1}$$

with a generic broken power-law spectrum given by

$$L_\nu \propto \begin{cases} \nu^{-0.5} & (1050 < \lambda < 1450\,\text{Å}), \\ \nu^{-1.5} & (\lambda < 1050\,\text{Å}). \end{cases} \tag{2}$$

This spectrum has an extreme UV (EUV) spectral index $\alpha_{\rm s} = 1.5$, consistent with radio-quiet quasars at lower redshift (Telfer et al. 2002). Note, however, the exact spectrum and hence ionizing luminosity of the quasar are rather uncertain at $z = 6$. Adopting a harder (softer) EUV spectral index will increase (decrease) the ionizing luminosity and the amount of photoheating around the quasar. We assume $M_{1450} = -27.4$ for the absolute magnitude of the quasar, corresponding to $\dot{N} = 2.74 \times 10^{57}\,{\rm s}^{-1}$ for our chosen spectrum. This magnitude corresponds to the M_{1450} obtained for SDSS J0818+1722 (Fan et al. 2006). Lastly, we use a duration of $t_{\rm Q} = 10^7$ yr for the phase when the quasar is optically bright (Haehnelt, Natarajan & Rees 1998; Martini 2004; Croton 2009). This parameter is also somewhat uncertain; we discuss the implications of this in more detail below.

In all our RT simulations, we initialize the ionization state of the IGM by assuming photoionization equilibrium with a UVB given by

$$J_\nu = J_{-21}\left(\frac{\nu}{\nu_{\rm H\,I}}\right)^{-3} \times \begin{cases} 1 & (\nu_{\rm H\,I} \le \nu < \nu_{\rm He\,II}), \\ 0 & (\nu_{\rm He\,II} \le \nu), \end{cases} \tag{3}$$

where J_{-21} is the amplitude of the UVB at the H I ionization threshold in proper units of 10^{-21} erg s^{-1} cm^{-2} Hz^{-1} sr^{-1}. We adopt a soft, stellar-like spectrum and assume H I and He I is already highly ionized by $z \le 6$ (e.g. Choudhury & Ferrara 2006; Becker et al. 2007; Pritchard, Loeb & Wyithe 2009). There is no evidence to suggest the hydrogen in the IGM surrounding SDSS J0818+1722 was significantly neutral prior to the quasar turning on (see Bolton & Haehnelt

[1] The ionization and thermal state of the IGM is fully recomputed in post-processing by our RT algorithm; we only use the hydrodynamical simulations for the initial conditions (gas density, peculiar velocity and temperature) of the gas. The gas overdensity and peculiar velocity field thus remain static in the RT calculation, and the hydrodynamical response of the gas to the photoheating induced by the quasar is not modelled. This effect should be small over the quasar lifetime we consider [$t_{\rm Q} \simeq 10^7$ yr, cf. the sound crossing time-scale $t_{\rm sc} \simeq 6.7 \times 10^8 (L/10\,\text{proper kpc})(T/10^4\,{\rm K})^{-1/2}$ yr and the dynamical time-scale $t_{\rm dyn} \simeq 4.2 \times 10^9 \Delta^{-1/2}\,[(1+z)/7]^{-3/2}$ yr].

2007b for a detailed discussion of this point). The exact choice of UVB spectral index at frequencies below the He II ionization threshold is unimportant, since the IGM temperature is initialized using the hydrodynamical simulations.

However, He II reionization, which must be driven by sources with hard ionizing spectra, is thought to be delayed until around $z \simeq 3$ when the number density of quasars begins to peak (Madau & Meiksin 1994; Furlanetto & Oh 2008). We model this delay by truncating the UVB above the He II ionization threshold. When the quasar turns on in our RT models, its hard ionizing spectrum will therefore reionize and photoheat the He II in the surrounding IGM, resulting in the so-called 'thermal proximity effect' (Miralda-Escudé & Rees 1994; Theuns et al. 2002b; Meiksin, Tittley & Brown 2010). The main effect of He II reionization by the quasar is to flatten the temperature–density relation and boost the temperature at mean density by ∼7000–9000 K within the proximity zone. The volume-weighted temperature–density plane from one of our RT simulation sets (model C) is displayed in Fig. 1. The final (approximate) temperature–density relation in the RT simulations after He II photoheating by the quasar is given in the last two columns of Table 1. As already discussed, the exact amount of heating depends on the assumed EUV spectral index. Note also that if the quasar lifetime is significantly shorter (longer) than the generic value of $t_Q = 10^7$ yr we have assumed, the extent of the region where He II is photoionized and heated by the quasar will be smaller (larger). In this work we consider the IGM within 33 Mpc of SDSS J0818+1722, a scale over which He II is fully ionized for

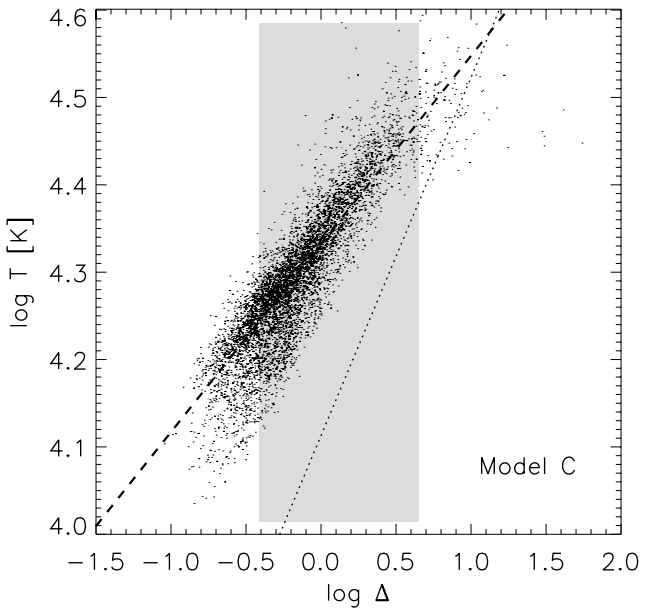

Figure 1. Scatter plot of the volume-weighted temperature–density plane for the 30 lines-of-sight used in model C. A random subset of pixels $250\,\mathrm{km\,s^{-1}} \leq v_H \leq 3250\,\mathrm{km\,s^{-1}}$ from the quasar redshift are displayed for clarity of presentation. The dashed line shows the final (approximate) power-law temperature–density relation after He II photoheating by the quasar, while the dotted line corresponds to the initial IGM temperature–density relation (taken from the hydrodynamical simulation) before the quasar turned on. There is significant scatter in the relationship between temperature and density following quasar activity. The shaded region corresponds to 95 per cent of all pixels with normalized flux $0.05 \leq F \leq 0.95$ in the corresponding proximity zone spectra. This gives an indication of the densities which transmission in the proximity zone is sensitive to (see the discussion in Section 4.2).

$t_Q = 10^7$ yr for our choice of quasar spectrum. Longer quasar lifetimes will make little difference to our results. However, if $t_Q \lesssim 10^7$ yr our models will overestimate the size of the He III region (and hence He II photoheating) around the quasar.

Lastly, the amplitude of the UVB at the H I ionization edge in equation (3), J_{-21}, is set using the density-dependent model presented in Wyithe et al. (2008). Briefly, this model accounts for the bias in the overdense region around the quasar host halo and is computed as a function of the proper time along the trajectory of a photon emitted by the quasar. This results in an enhanced contribution to the UVB near the quasar host halo relative to the mean IGM. The latter is calibrated to match measurements of the H I photoionization rate derived from the observed Lyα forest opacity at $4 \leq z \leq 6$ (Bolton & Haehnelt 2007c). A detailed description of the model and its application to high-redshift Lyα forest spectra may be found in Wyithe et al. (2008). Note that the UVB amplitude is important for the observed *size* of a proximity zone (Wyithe et al. 2008), but the precise value of this parameter is not important for the gas temperatures around the quasar we explore in this work.

2.4 Spectra construction

We follow a standard procedure to construct synthetic proximity zone spectra from the output of each of our RT models (e.g. Theuns et al. 1998). The spectra are then convolved with a Gaussian instrument profile with FWHM $= 6.7\,\mathrm{km\,s^{-1}}$, resampled on to pixels of width $2.1\,\mathrm{km\,s^{-1}}$, and Gaussian distributed noise with S/N = 12 is added. These values are chosen to match the observed spectrum of SDSS J0818+1722. To account for uncertainties in the continuum placement on the observed data, we renormalize the synthetic data by the highest flux in $1000\,\mathrm{km\,s^{-1}}$ segments along each line-of-sight. Although this will not account for continuum uncertainties perfectly, it will approximate the way the continuum is fitted to the real data and should reduce a potential systematic bias. We consider this issue further in Section 4.2. Lastly, we rescale the optical depths in each pixel of the synthetic spectra by a constant to match the mean flux of the observed data, $\langle F \rangle_{\rm obs} = I_{\rm obs}/I_{\rm cont} = 0.486$, within the range $250\,\mathrm{km\,s^{-1}} \leq v_H \leq 3250\,\mathrm{km\,s^{-1}}$(2.6 Mpc $\leq R \leq$ 33.3 Mpc). An example spectrum drawn from model C is displayed in Fig. 2 along with the observed line-of-sight we use in this work, SDSS J0818+1722.

3 ANALYSIS PROCEDURE

3.1 Voigt profile fitting

In this work, we use the cumulative probability distribution function (CPDF) of the Doppler widths of Lyα absorption lines, b, in the proximity zone as our probe of the IGM temperature. The CPDF has the advantage of fully using the limited data available (our analysis of SDSS J0818+1722 yields only 25 Doppler widths) and it avoids binning and the associated loss of information. Although the absorption lines will not all be purely thermally broadened (Theuns, Schaye & Haehnelt 2000), all absorption lines will nevertheless be smoothed on a scale associated with the thermal broadening kernel. The entire CPDF is therefore sensitive to the IGM temperature. However, the lack of a direct relationship between the Doppler parameters and temperature means detailed synthetic spectra are crucial for calibrating this statistic.

We obtain the CPDF for both the observed and synthetic data via Voigt profile fitting. We perform the analysis using an automated

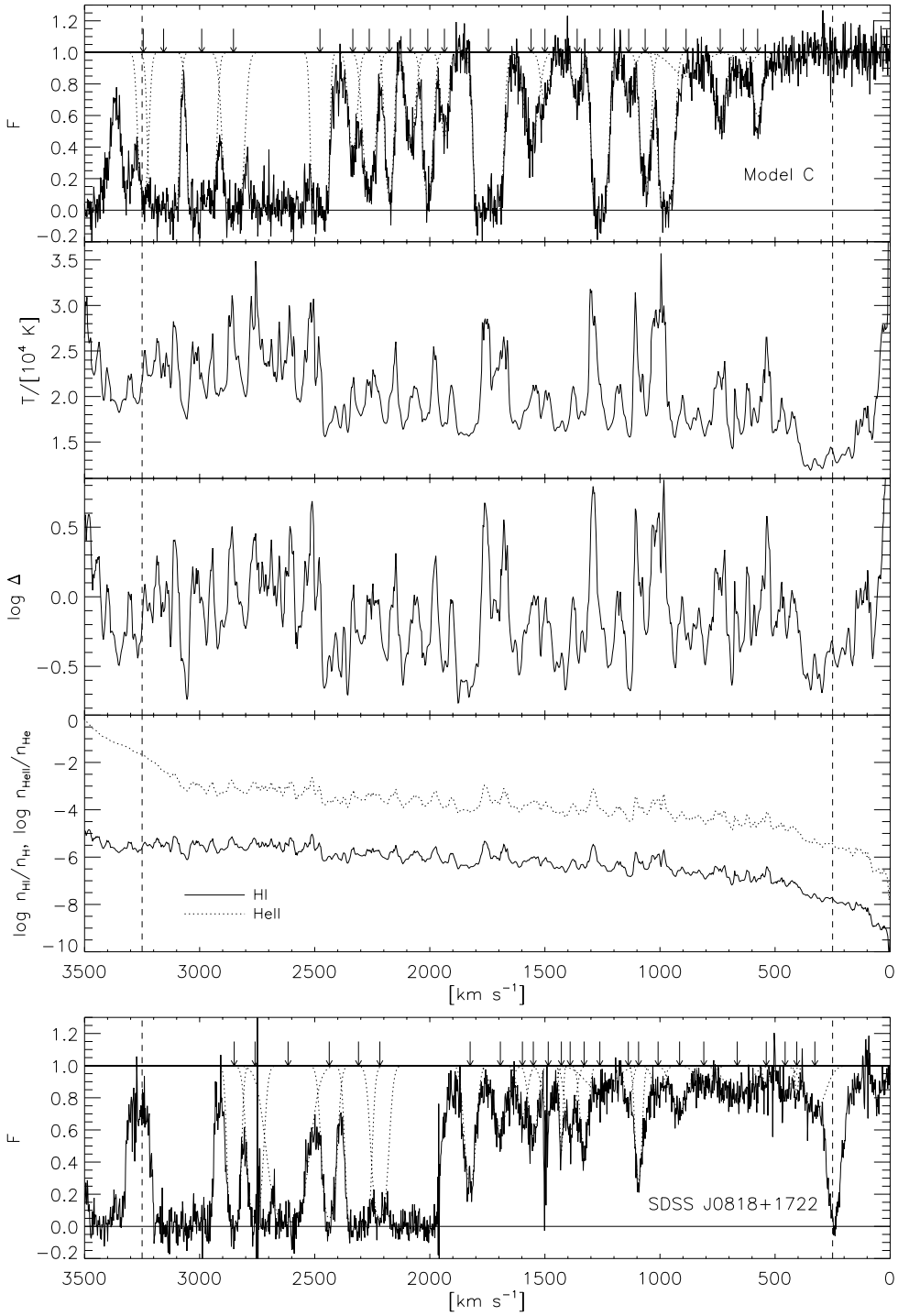

Figure 2. Upper four panels: the topmost panel shows an example synthetic spectrum at $z = 6$ drawn from model C. The quasar is situated at $v_{\rm H} = 0\,{\rm km\,s^{-1}}$ on the right-hand side of each panel. The vertical dashed lines indicate the velocity range over which Voigt profiles are fitted to the data, $250\,{\rm km\,s^{-1}} \leq v_{\rm H} \leq 3250\,{\rm km\,s^{-1}}$ (2.6 Mpc $\leq R \leq$ 33.3 Mpc). The vertical arrows mark the redshift of the Voigt profile fits made with VPFIT, while the fits themselves are shown by the dotted curves. The corresponding gas temperature, overdensity and fractional abundances of H I and He II along the line-of-sight are shown in the subsequent three panels. Lower panel: Keck/HIRES spectrum of SDSS J0818+1722 at $z = 6.00$.

version of the Voigt profile fitting package VPFIT.[2] We choose to include lines in the range $250\,{\rm km\,s^{-1}} \leq v_{\rm H} \leq 3250\,{\rm km\,s^{-1}}$ (2.6 Mpc $\leq R \leq$ 33.3 Mpc) from the quasar redshift only. The lower limit is chosen to avoid edge effects in the synthetic data, while beyond $3250\,{\rm km\,s^{-1}}$ it becomes impossible to fit Voigt profiles reliably due to the increasing opacity of the IGM. All lines with relative errors in excess of 50 per cent are rejected from our final analysis (Schaye et al. 1999). This removes most of the very narrow lines which tend to be in blends. We also discard all lines with Doppler parameters

[2] http://www.ast.cam.ac.uk/∼rfc/vpfit.html

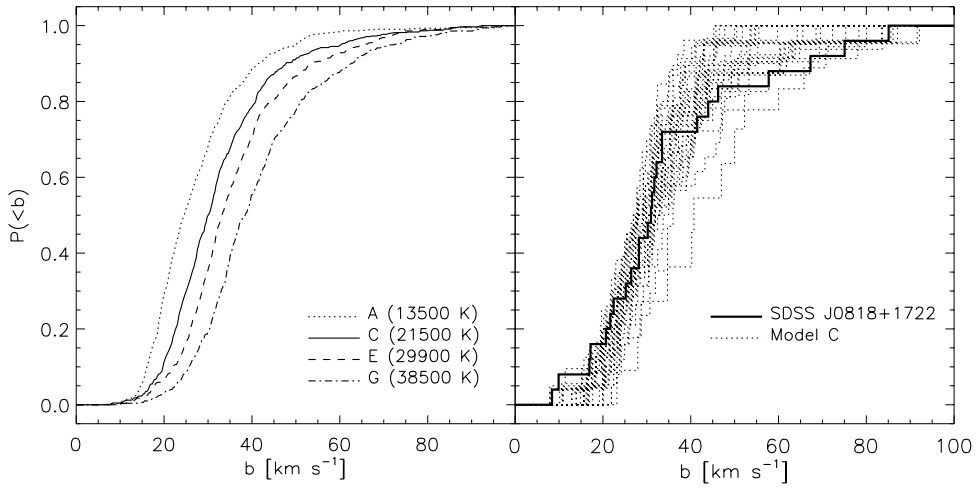

Figure 3. Left-hand panel: the Doppler parameter CPDF for four of our RT models. The simulated CPDFs correspond to the data from all 30 lines-of-sight in each model. The CPDF is clearly sensitive to the gas temperature in the quasar proximity zone, with the CPDF shifting to higher b-values for increasing temperature. The median Doppler parameters for the simulated models displayed are $b_{\rm med} = (24.5, 30.0, 33.2, 38.4)\,{\rm km\,s^{-1}}$. Right-hand panel: the observed distribution for SDSS J0818+1722 (thick solid curve), which has $b_{\rm med} = 31.0\,{\rm km\,s^{-1}}$, in comparison to the CPDFs for the individual lines-of-sight in model C. This gives an indication of the scatter in the CPDF from one line-of-sight to the next.

$b > 100\,{\rm km\,s^{-1}}$ and column densities $\log N_{\rm HI} > 17$. Broad shallow lines with $b > 100\,{\rm km\,s^{-1}}$ are almost always added near the continuum level to improve the overall fit, while strong lines with damping wings tend to be added near the edge of the proximity zone where there is very little transmission. These cuts remove 32 and 25 per cent of identified Voigt profiles in the observed and simulated data sets, respectively. Lastly, we note that some ambiguous fits still remain in our analysis. However, since we treat the observed and synthetic spectra in exactly the same way we expect any potential bias to be small.

The sensitivity of the resulting Doppler width CPDF to the gas temperature is demonstrated in the left-hand panel of Fig. 3, where the CPDFs for four of our RT simulation sets are displayed. The CPDFs correspond to the data from all 30 lines-of-sight in each RT simulation set. It is clear that the entire CPDF is sensitive to the gas temperature in the proximity zone. Increasing the temperature of the IGM shifts the CPDF towards higher b-parameter values, increasing the median Doppler parameter. The median Doppler parameters for the models displayed, which span a range of $13\,500\,{\rm K} \leq T_0 \leq 38\,500\,{\rm K}$, are $b_{\rm med} = (24.5, 30.0, 33.2, 38.4)\,{\rm km\,s^{-1}}$. The observed distribution, which has $b_{\rm med} = 31.0\,{\rm km\,s^{-1}}$, is displayed as the thick solid curve in the right-hand panel of Fig. 3 along with the CPDFs for the *individual* lines-of-sight in model C. This gives an indication of the scatter in the CPDF from one line-of-sight to the next for the same underlying model. A two-distribution Kolmogorov–Smirnov (KS) test (Press et al. 1992) using the average (all 30 lines-of-sight) and observed CPDF yields $P_{\rm KS} = 0.06, 0.94, 0.21$ and 4×10^{-4} for models A, C, E and G, respectively. This does not rule out the null hypothesis that the observed CPDF is drawn from the same distribution as models A, C and E.

3.2 Monte Carlo analysis

We now turn to describing the methodology which lies at the heart of our analysis procedure. We use our synthetic spectra to construct Monte Carlo realizations of the Doppler parameter CPDF for a range of models with different IGM temperatures. To quantify the amount of scatter in the simulated CPDFs, we use a 'D-statistic' that is very similar to the parameter used in a KS test. For each line-of-sight in a given simulation set, the D-statistic is the maximum difference between the Doppler parameter CPDF for that line-of-sight and the CPDF for all 30 lines-of-sight, such that

$$D_{\rm i} = \max|P(<b)_{\rm i} - P(<b)_{\rm all}|,\ i = 1, \ldots, 30, \tag{4}$$

where we preserve the sign of the difference. The D-statistic CPDF for a model with known temperature T_0, $P(<D|T_0)$, can then be constructed. If the D-statistic for an observed line-of-sight, $D_{\rm obs}$, is known, the D-statistic CPDF may then be used in combination with the Bayes theorem to infer a confidence interval for the *observed* T_0. The cumulative probability for the observed temperature at mean density T_0 given $D_{\rm obs}$ is

$$P(<T_0|D_{\rm obs}) = \int_0^{T_0} \frac{{\rm d}P(<T_0'|D_{\rm obs})}{{\rm d}T_0'}{\rm d}T_0'. \tag{5}$$

An application of Bayes theorem, $p(T_0|D_{\rm obs}) \propto p(D_{\rm obs}|T_0)p(T_0)$, then leads to

$$P(<T_0|D_{\rm obs}) = K\int_0^{T_0} \frac{{\rm d}P(<D_{\rm obs}|T_0')}{{\rm d}D}\frac{{\rm d}P(<T_0')}{{\rm d}T_0'}{\rm d}T_0', \tag{6}$$

where K is a constant which normalizes the total probability to unity and $p(T_0)$ is the prior on T_0. In order to obtain our constraint on T_0, we evaluate the derivative of the D-statistic CPDF derived from our simulations at $D_{\rm obs}$ in equation (6). We adopt a flat prior, $p(T_0)$, over the range of temperatures explored in the simulations. These are set by the combined effect of the initial temperature of the hydrodynamical simulations and the subsequent heat input from He II photoheating by the quasar, such that $13\,500\,{\rm K} \leq T_0 \leq 38\,500\,{\rm K}$. This is intended to represent a reasonable range for the IGM temperature following H I and He II reionization, and is consistent with temperatures predicted in simulations of He II reionization (McQuinn et al. 2009).

The main drawback of our method is the computational expense of constructing and analyzing synthetic spectra. We have seven hydrodynamical simulations with different values of T_0, so we may only evaluate $\frac{{\rm d}P(<D_{\rm obs}|T_0)}{{\rm d}D}$ at seven discrete points. We use a cubic spline to interpolate between these to obtain a continuous distribution for integration. With the current method, finer sampling of

the distribution would require running and analyzing additional simulations. A less expensive approach would be to use a single hydrodynamical simulation and simply impose a range of initial temperature–density relations on the IGM. However, this would decouple the initial gas temperature from the gas hydrodynamics and incorrectly model the effect of pressure (Jeans) smoothing on the gas distribution (e.g. Pawlik, Schaye & van Scherpenzeel 2009; Peeples et al. 2009). Furthermore, there are only 30 line-of-sight in each simulation set, which makes the derivative of the D-statistic CPDF somewhat noisy. This could be remedied by analyzing more lines-of-sight, although the Voigt profile analysis procedure is time consuming. For this first attempt we have instead elected to perform a easily manageable number of simulations. We therefore smooth the D-statistic CPDF with a Gaussian filter of width $\sigma = 0.025$ before computing its derivative. We find this to be the optimum smoothing width, and we have verified that alternative choices of $\sigma = 0.01$ and 0.05 do not significantly alter our final results.

On the other hand, the major advantage of our approach is that it avoids binning the data and is non-parametric. However, our method implicitly assumes that the synthetic spectra accurately represent the observed data. If our synthetic models are significantly in error, any temperature constraints will be unreliable. For this reason, we now turn to consider numerical convergence and some of the possible systematic errors which may affect our analysis.

4 CONVERGENCE AND SYSTEMATICS

4.1 Numerical convergence

We first consider the effect of mass resolution and box size on the simulated Doppler width CPDF in Fig. 4. The left-hand panel displays the impact of mass resolution on the CPDF, while the right-hand panel indicates the effect of box size. The CPDF is well converged for our fiducial mass resolution ($9.2 \times 10^4\,h^{-1}\,\mathrm{M}_\odot$), although note our lowest resolution model (L1, $5.9 \times 10^6\,h^{-1}\,\mathrm{M}_\odot$) does not correctly resolve the linewidths. It is also encouraging that the box size, and hence the amount of large-scale power in our simulations, has little effect on the CPDF. The maximum halo mass in the $40\,h^{-1}$ Mpc box is $4.33 \times 10^{11}\,h^{-1}\,\mathrm{M}_\odot$, which is an order of magnitude greater than the largest halo in our fiducial simulation volume of $10\,h^{-1}$ Mpc ($6.13 \times 10^{10}\,h^{-1}\,\mathrm{M}_\odot$). However, this has little effect on the CPDF because we are measuring Doppler parameters in the ambient IGM rather than in the host environment of the quasar. Note, however, that the maximum halo mass we consider is still smaller than the expected host masses of $\sim 10^{12}$–$10^{13}\,\mathrm{M}_\odot$ for bright $z \simeq 6$ quasars (Walter et al. 2004; Volonteri & Rees 2006; Li et al. 2007). We nevertheless find it is much more important to resolve absorption linewidths correctly in our models using a high-resolution simulation rather than using a larger volume. We conclude that our simulations are well converged with box size and mass resolution for the purpose of this study; any differences in the CPDF are small compared to the contrast between models with different temperatures.

4.2 Systematic uncertainties

4.2.1 The mean flux

The effect of four different systematic uncertainties on the simulated Doppler parameter CPDF is displayed in Fig. 5. The first we consider is the mean transmitted flux, $\langle F \rangle = \langle e^{-\tau_i} \rangle$, where τ_i is the optical depth in pixel i. As discussed earlier, when constructing our synthetic spectra we follow a standard procedure and rescale the H I photoionization rate in post-processing by a fixed constant to produce spectra which match the $\langle F \rangle$ of the observed data, $\langle F \rangle = 0.486$. It is therefore important to assess how a different value of $\langle F \rangle$ may influence the Doppler parameter CPDF.

The effect of the mean flux in the region of the spectra where lines are fitted, $250\,\mathrm{km\,s^{-1}} \leq v_\mathrm{H} \leq 3250\,\mathrm{km\,s^{-1}}$, is shown in the upper-left panel of Fig. 5. We perform a Voigt profile analysis on two further models which are identical to model C, except from being rescaled in post-processing to have $\langle F \rangle = 0.386$ and $\langle F \rangle = 0.586$. It is clear that the mean flux has little effect on the Doppler parameter CPDF, changing the median Doppler parameter by less than $0.5\,\mathrm{km\,s^{-1}}$ for the examples shown. This is expected theoretically; as pointed out by Hui & Rutledge (1999), the ionizing background does not play a role in setting the *shape* of an absorption line – and hence its Doppler parameter – around the peak optical depth. Note, however, that the mean flux will indeed be important for the line column densities, decreasing N_HI at a fixed density when the

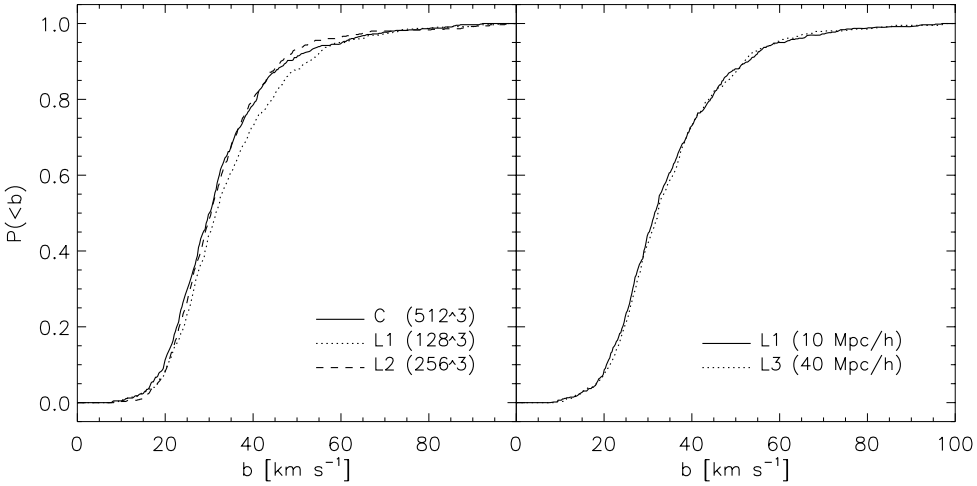

Figure 4. Numerical convergence tests for the simulated Doppler width CPDF. The CPDFs correspond to the data from all 30 lines-of-sight in each RT simulation set. Left-hand panel: test of the hydrodynamical simulation mass resolution. Although the CPDF from the lowest resolution L1 model has not converged, the distribution is well captured for our fiducial mass resolution of $9.2 \times 10^4\,h^{-1}\,\mathrm{M}_\odot$ (model C). Right-hand panel: test of the simulation box size. The CPDF is well converged, indicating our choice of a $10\,h^{-1}$ Mpc box is sufficient.

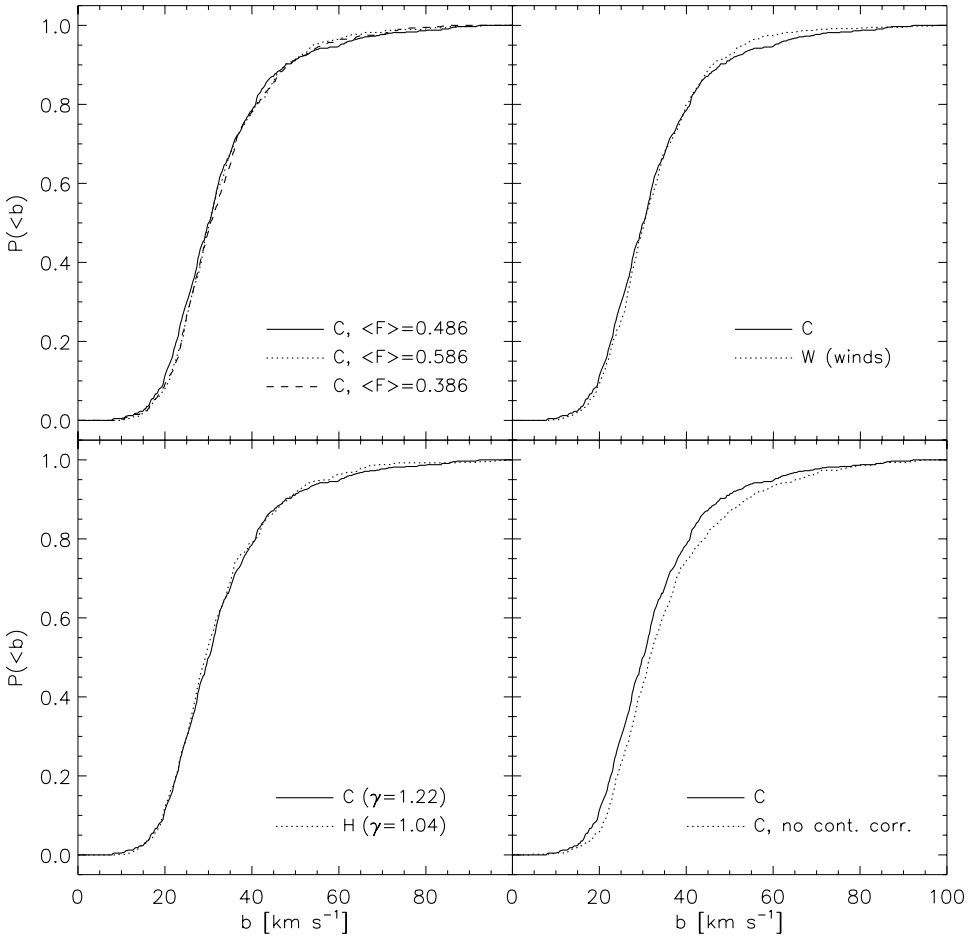

Figure 5. Systematic uncertainties and their effect on the Doppler parameter CPDF. Upper left: The effect of the mean flux measured over $250\,\mathrm{km\,s^{-1}} \leq v_\mathrm{H} \leq 3250\,\mathrm{km\,s^{-1}}$. The fiducial value in this work is $\langle F \rangle = 0.486$ (solid curve), corresponding to value measured from the spectrum of SDSS J0818+1722. Two further models with $\langle F \rangle = 0.386$ (dashed curve) and $\langle F \rangle = 0.586$ (dotted curve) are displayed for comparison. Upper right: the impact of feedback in the form of strong galactic winds, $v_\mathrm{w} = 448\,\mathrm{km\,s^{-1}}$, implemented using the multi-phase star formation model of Springel & Hernquist (2003). Model C (solid curve) has no winds while model W (dotted line) is identical to model C in all respects besides the inclusion of galactic winds. Lower left: the effect of the temperature–density relation slope, γ. Model C (solid curve) has $\gamma \sim 1.2$ while model H (dotted curve) has $\gamma \sim 1$ following He II photoheating by the quasar. Lower right: the effect of uncertainty in the continuum placement. The dotted curve shows the CPDF obtained before the standard continuum correction is applied to the synthetic data (see the main text for details), while the solid curve is obtained from the data after correction. Lowering the continuum (it is never raised in the correction) always reduces the median Doppler parameter.

photoionization rate is raised. We find that uncertainties in the mean flux should not be important for this work.

4.2.2 Galactic winds

We reach a similar conclusion regarding the impact of galactic winds on the CPDF in the upper-right panel of Fig. 5. We have constructed spectra from model W. This uses the multi-phase star formation and winds model of Springel & Hernquist (2003), but in all other respects is identical to model C. The wind velocity is $v_\mathrm{w} = 448\,\mathrm{km\,s^{-1}}$, which is an extreme model in the sense that it assumes all energy produced by supernovae is converted into kinetic energy. The differences in the CPDF are again small, with $b_\mathrm{med} = 30.0\,\mathrm{km\,s^{-1}}$ and $30.3\,\mathrm{km\,s^{-1}}$ for the C and W models, respectively.

4.2.3 The temperature–density relation

We do not attempt to measure the density dependence of the IGM temperature in this work; this is difficult to achieve for a single quasar spectrum at $z = 6$. Instead, we focus on constraining the temperature at a mean density, T_0. The assumed power-law slope of the temperature–density relation, γ, is therefore a systematic uncertainty in our analysis. Note again that a single power-law temperature–density relation, $T = T_0 \Delta^{\gamma-1}$, does not provide an accurate description of the IGM thermal state in our simulations (e.g. Fig. 1, and see also Bolton, Meiksin & White 2004; Trac, Cen & Loeb 2008). Nevertheless, for the present we find it convenient to characterize the relationship between temperature and density as scatter around a power law.

In the lower-left panel of Fig. 5 we demonstrate the effect of γ on the Doppler parameter CPDF. Models C and H have $\gamma \sim 1.2$ and $\gamma \sim 1.0$, respectively, with almost identical temperatures at mean density. Interestingly, there is very little difference between the CPDFs, with $b_\mathrm{med} = 30.0\,\mathrm{km\,s^{-1}}$ and $29.3\,\mathrm{km\,s^{-1}}$. This is the result one would expect if the CPDF is sensitive to gas densities around the cosmic mean at $z = 6$. We may test if this is the case by evaluating the range of optical depth weighted overdensities, $\Delta_\tau = \sum \Delta_\mathrm{i} \tau_\mathrm{i} / \sum \tau_\mathrm{i}$ (e.g. Schaye et al. 1999), over the range $250\,\mathrm{km\,s^{-1}} \leq v_\mathrm{H} \leq 3250\,\mathrm{km\,s^{-1}}$ in our simulated spectra. In Fig. 1 the shaded region corresponds to the range of Δ_τ including 95 per cent of the

pixels with $0.05 \leq F \leq 0.95$ in the proximity zone: $0.4 \leq \Delta_\tau \leq 4.4$, with a median value of $\Delta_\tau = 0.95$. The transmission in the proximity zone thus predominantly probes the IGM around mean density, and it is clear why different values of γ have only a small impact on the CPDF. Future studies at $z \simeq 6$ could utilize higher-order Lyman series absorption sensitive to higher gas densities to obtain constraints on γ (Dijkstra, Lidz & Hui 2004; Furlanetto & Oh 2009). For now, based on these results we shall assume the impact of this parameter on our analysis is small.

4.2.4 *Continuum placement*

We now consider the effect of the uncertainty in the intrinsic emission of the quasar, $I_{\rm cont}$, where $I_{\rm obs} = I_{\rm cont} e^{-\tau}$. At $z = 2$–3, the typical uncertainty in the continuum placement in high-resolution Lyα forest data is estimated to be 1–4 per cent (Tytler et al. 2004; Kirkman et al. 2005; Kim et al. 2007; Faucher-Giguère et al. 2008). However, towards higher redshift, the increasing opacity of the IGM means the transmitted flux recovers to the unabsorbed level less and less frequently. Identifying the continuum level accurately becomes very challenging, with an estimated uncertainty of up to $\sim$20 per cent even for the high-quality spectrum considered here. Continuum estimation therefore presents a significant systematic uncertainty in our analysis.

As it is difficult to identify where the continuum lies on the observed data at $z \simeq 6$, our approach in this work is instead to treat the simulated spectra in the same way as the observational data. The observational data were normalized using a slowly varying spline fit, adjusted by hand to roughly follow the profile of the transmission peaks in the proximity zone. The continuum correction we apply to our data, described in Section 2.4, is designed to mimic this. To recap, we renormalize the synthetic data by the highest flux in $\sim$1000 km s^{-1} segments along each line-of-sight. Although this means we may not correctly identify the 'true' continuum, in principle this procedure should at least reduce any bias (see Rauch et al. 1997; McDonald et al. 2000 for similar approaches to this problem at lower redshift).

The lower-right panel of Fig. 5 displays the effect of the continuum correction on the CPDF, which always acts to *lower* the continuum on the synthetic spectra. Although this correction has little effect near the redshift of the quasar – the IGM here is very highly ionized and the flux regularly reaches the unabsorbed level in our simulations – it becomes larger towards the blue-most edge of the region we fit lines to; the Lyα transmission rarely reaches $F = 1$ beyond $\simeq$2000 km s^{-1}. On average the maximum correction is around 10 per cent, but it can be up to 25 per cent for some lines-of-sight at the edge of the proximity zone. The correction shifts the CPDF to lower Doppler widths, effectively narrowing the absorption lines as the continuum is lowered. For reference, the median Doppler parameters are $b_{\rm med} = 30.0$ km s^{-1} (with correction) and 31.6 km s^{-1} (no correction). Consequently, if the continuum on the *observed* data has been placed high (the correction always lowers the continuum) relative to the corrected synthetic data, we will infer gas temperatures which are $\sim$2000 K too high. This is clearly an important systematic to consider when the statistical error bars approach the level of a few thousand degrees.

4.2.5 *Metals*

Lastly, we consider the impact of misidentified metal lines on the Doppler parameter CPDF. A full treatment would require modelling the distribution and abundance of metals in the IGM at $z = 6$, which is beyond the scope of this work. However, assuming metal lines are typically narrower than the Lyα lines, we may estimate the effect of any metal contamination as follows. In our spectrum of SDSS J0818+1722, there are roughly two metal absorption lines per 3000 km s^{-1} interval redwards of the Lyα emission line. Over the same interval in the proximity zone, by comparison, we include 25 lines in our Doppler parameter CPDF. We may therefore estimate the effect of possible metal contamination by excluding the two narrowest lines identified in the proximity zone of SDSS J0818+1722 from our analysis, $b = 8.6$ and 10.0 km s^{-1} (cf. the narrowest line in model C, $b = 7.9$ km s^{-1}). This changes the median Doppler parameter for SDSS J0818+1722 from $b_{\rm med} = 31.0$ to 31.3 km s^{-1}. We estimate that metal contamination may then at most systematically *decrease* any constraint on the IGM temperature in the proximity zone by $\sim$2000 K. As we shall see, the systematic uncertainties discussed in this section are small compared to the statistical uncertainty for an individual line-of-sight. Consequently, we do not consider these in our final analysis. However, we note that future analyses with smaller statistical errors will require a more detailed treatment of continuum placement and metal contamination.

5 RESULTS

5.1 Test of the analysis procedure

We are now ready to progress to the main result of this paper. However, before proceeding further we briefly test our methodology. We apply the analysis procedure described in Section 3 to two synthetic lines-of-sight for which we already know the 'true' temperature. In this way we may check if our procedure correctly recovers the temperature of the IGM around a quasar.

The test of our methodology is displayed in Fig. 6. Each panel shows the constraint on T_0 for a synthetic line-of-sight in the form of a cumulative distribution for T_0, $P(< T_0|D_{\rm obs})$, computed using equation (6). The lines-of-sight are drawn from the B and F models which have $T_0^{\rm true} = 17\,500$ and 34 400 K, respectively. These temperatures are represented by the filled circles in Fig. 6. The shaded range in both panels display the 68 per cent confidence intervals around the inferred median value of T_0. Due to the limited statistical power of a single line-of-sight the uncertainties on the constraints are large, $T_0 = 18\,900 \pm ^{4400}_{6900}$ K ($\pm^{12700}_{5000}$ K) and 32 400 $\pm ^{4100}_{6000}$ K ($\pm^{5600}_{10100}$ K) at 68 (95) per cent confidence around the median. Nevertheless, it is encouraging that our procedure correctly distinguishes between models with different temperatures.

5.2 Measurement of T_0 for SDSS J0818+1722

Fig. 7 displays the cumulative distribution for $T_0, P(< T_0|D_{\rm obs})$, derived for the line-of-sight SDSS J0818+1722 at $z = 6$. We find a median temperature of $T_0 = 23\,600 \pm ^{5000}_{6900}$ K ($\pm^{9200}_{9300}$ K), which lies in the middle of our simulation grid. The CPDF flattens towards high temperatures, indicating that temperatures of $T_0 \gtrsim 33\,000$ K are inconsistent with the observed Doppler parameter data. The lower limit on T_0 is influenced by our choice of prior for T_0 following H I and He II reionization, $13\,500\,{\rm K} \leq T_0 \leq 38\,500$ K (e.g. McQuinn et al. 2009). However, the gradient of the CPDF does begin to flatten towards the lower edge of the simulation grid, suggesting the probability of temperatures below 13 500 K is fairly low.

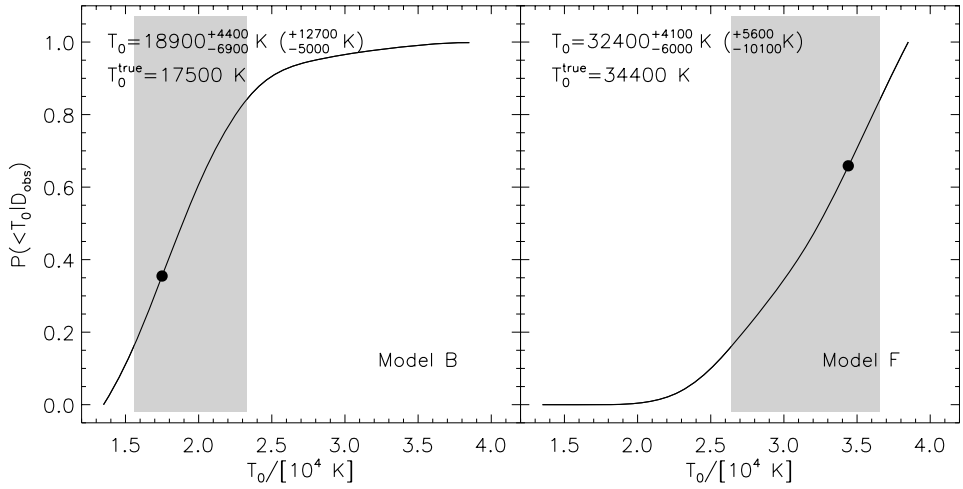

Figure 6. Test of our analysis procedure using two random lines-of-sight drawn from our simulated data set. The lines-of-sight in the left- and right-hand panels are drawn from models B and F, respectively. The solid curves show the cumulative distribution for $T_0, P(<T_0|D_{\rm obs})$, computed using equation (6). The filled circles show the 'true' temperature of the models at mean density, while the shaded regions indicate the 68 per cent confidence interval around the recovered median. The numerical values in the top left of each panel correspond to the recovered median temperature and the 68 (95) per cent confidence intervals. The large uncertainties are dominated by the statistical error associated with measuring the temperature along a single line-of-sight.

5.3 Consistency check using proximity zone sizes

In an analysis of observed proximity[3] zone sizes, WBH08 found the *absolute* sizes of the proximity zones in their RT models were systematically smaller than the observational data. Based on their RT modelling of the proximity zones WBH08 further found that the size of the proximity zones should depend primarily on the background H I photoionization rate and the temperature of the IGM, where the latter impacts on the ionization state of hydrogen through the H II recombination coefficient, $\alpha_{\rm HII} \propto T^{-0.7}$. WBH08 therefore suggested that the IGM temperature in quasar proximity zones at $z \simeq 6$ may be as high as $T_0 \simeq 40\,000$ K, perhaps resulting from the recent reionization of H I and He II by a hard ionizing spectrum. Imposing a high temperature of $T_0 \simeq 40\,000$ K was found to resolve the discrepancy between the sizes of observed and simulated proximity zones. Our results, albeit for a single quasar proximity zone, are inconsistent with this interpretation. If our constraint is representative of the other $z \simeq 6$ lines-of-sight, an alternative explanation would be required to explain the proximity zone sizes.

In this paper, we revisit the question of proximity zone size using our new suite of high-resolution simulations. This analysis provides a consistency check of our temperature constraints. We have repeated the proximity zone size analysis performed by WBH08 (see their section 5 for details) using our model C simulations. This model has $T_0 = 21\,500$ K, similar to the median value derived for SDSS J0818+1722. The main difference between our work and WBH08 is the much higher resolution of our hydrodynamical simulations. WBH08 used a simulation with gas particle masses of $4.3 \times 10^7\ h^{-1}\,{\rm M}_\odot$ and a box size of 60 h^{-1} Mpc. This rather large volume and low resolution was chosen so that the largest halo in the simulation volume was reasonably massive, with a mass of $2.7 \times 10^{12}\ h^{-1}\,{\rm M}_\odot$. Here, we use much higher resolution hydrodynamical simulations which resolve the low-density IGM (Bolton & Becker 2009). In this work, we have demonstrated that mass resolution is more important than the halo mass for correctly resolving transmission from the Lyα forest in the quasar proximity zone. We also now use the observational data compiled by Carilli et al. (2010), which include seven additional quasars and an improved determination of systemic redshifts.

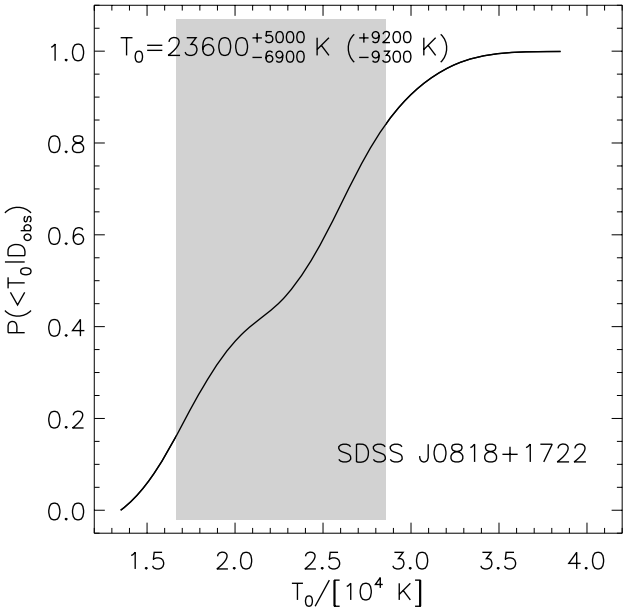

Figure 7. Constraint on the temperature of the IGM at mean density at $z = 6$ in the proximity zone of the quasar SDSS J0818+1722. The solid curves show the cumulative distribution for $T_0, P(<T_0|D_{\rm obs})$, obtained using equation (6), while the shaded region indicates the 68 per cent confidence interval around the median. The numerical value in the top left of the panel corresponds to the recovered median temperature at mean density and the 68 (95) per cent confidence intervals.

[3] WBH08 refer to these regions as highly ionized 'near-zones', reflecting the fact that it can be difficult to determine whether these transmission windows are due to H II regions embedded in a significantly neutral IGM or a region of enhanced ionization (i.e. the proximity effect) in an otherwise highly ionized IGM at $z > 6$ (Bolton & Haehnelt 2007b; Maselli et al. 2007; Lidz et al. 2007). However, in this work we mainly concentrate on the ionized IGM at $z \leq 6$, and so for simplicity we refer to these regions as proximity zones throughout.

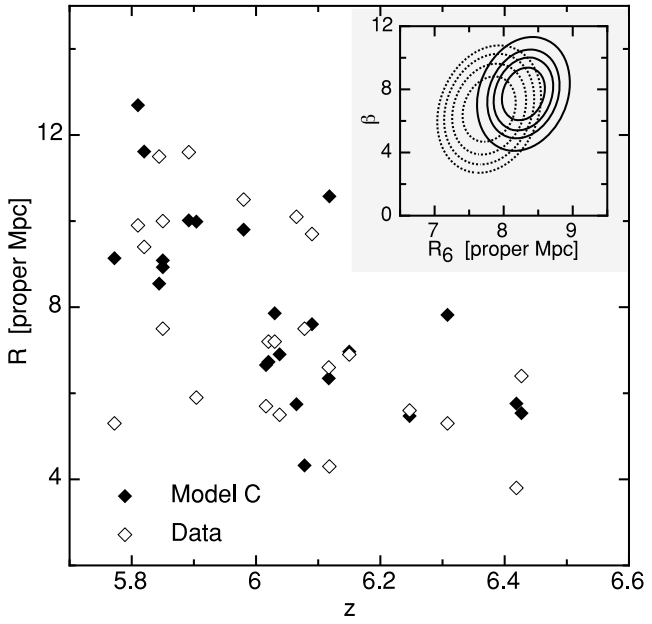

Figure 8. Proximity zone sizes (proper Mpc) measured from synthetic spectra (filled diamonds) compared to observational data taken from Carilli et al. (2010). The simulated spectra are constructed from model C, which has gas temperatures consistent with our constraint from the proximity zone of SDSS J0818+1722. Following Fan et al. (2006) and Carilli et al. (2010), the sizes have been rescaled to a common AB magnitude of $M_{1450} = -27$ by assuming the proximity zone size is proportional to $\dot{N}^{1/3}$. The inset displays the best-fitting linear relation to the observed and synthetic data, parametrized as $R = R_6 - \beta(z - 6)$. The dotted and solid contours represent the 68, 84, 91 and 97 per cent bounds on single parameters for the observed and simulated data, respectively.

The results are displayed in Fig. 8. The filled diamonds correspond to the simulation data, while the open diamonds show the observational data compiled in Carilli et al. (2010). Following Fan et al. (2006) and Carilli et al. (2010), the sizes[4] have been rescaled to a common AB magnitude of $M_{1450} = -27$ by assuming the proximity zone size is proportional to $\dot{N}^{1/3}$, where $\dot{N}$ is defined by equation (1). Since our hydrodynamical simulation outputs are all at $z = 6.01$, we must also rescale the physical gas density in the simulations by $(1 + z)^3$ to match the redshifts of the quasars, $5.77 \leq z \leq 6.43$. The inset shows the constraints on the amplitude and slope of the proximity zone evolution using the parametrization $R = R_6 - \beta(z - 6)$, where R_6 is the value of the proximity zone size at $z = 6$ and β is the slope of the evolution. The dotted and solid contours represent the 68, 84, 91 and 97 per cent bounds on single parameters for the observed and simulated data, respectively. We find that the simulated data are in reasonable agreement with the observational data without having to arbitrarily impose high gas temperatures as in WBH08. This is because the much higher mass resolution of our hydrodynamical simulation increases the transmission from the underdense regions. These dominate the transmission towards the edge of the proximity zone, increasing the extent of these regions. We have also verified that $T_0 > 33\,000$ K produces proximity zone sizes which are too large by performing the same analysis using model G. Our temperature constraint from SDSS J0818+1722 is therefore consistent with the rapid size evolution of proximity zones at $z \simeq 6$, providing a useful check of our modelling and results. As already noted by WBH08, the rapid evolution in the sizes of observed proximity zones also suggests that the IGM in the environment of bright quasars is already highly ionized by $z \simeq 6$.

[4] Following the observational definition of Fan et al. (2006), WBH08 define the extent of the proximity zones as the distance from the quasar at which the transmission falls below 10 per cent after smoothing by a top-hat filter of width 20 Å.

6 IMPLICATIONS FOR THE REIONIZATION HISTORY AROUND SDSS J0818+1722

This work has focused on measuring the IGM temperature within close proximity to a quasar. Although this means we are unable to place constraints on the thermal state of the 'average' IGM as previous studies have done at lower redshift (e.g. Schaye et al. 2000; Ricotti et al. 2000; McDonald et al. 2001; Zaldarriaga 2002; Lidz et al. 2009), it does enable us to study the IGM thermal history in the quasar's immediate vicinity (Miralda-Escudé & Rees 1994; Theuns et al. 2002b; Meiksin et al. 2010). The proximity zone temperature will encode information on when the IGM near SDSS J0818+1722 was first reionized. The long time-scale associated with adiabatic cooling in the low-density IGM enables the gas to retain a 'fossil record' of its initial temperature following reionization (Hui & Gnedin 1997; Theuns et al. 2002a; Hui & Haiman 2003). Adopting an approach similar to that employed by Hui & Haiman (2003), we may therefore use our temperature measurement to place an *upper limit* on the possible redshift of H I reionization around SDSS J0818+1722. We pose the following question: 'what is the maximum amount of time available for the IGM to cool following H I reionization before the temperature becomes inconsistent with our T_0 constraint at $z = 6$?'

To address this question, we shall compute the thermal evolution of a gas parcel at mean density in the presence of a power-law ionizing spectrum (e.g. Hui & Haiman 2003; Furlanetto & Oh 2009). Our code follows photoionization and heating, collisional ionization, radiative cooling, Compton cooling and adiabatic cooling for six species (H I, H II, He I, He II, He III, e^-). We use the rates compiled in Bolton & Haehnelt (2007b) with the exception of the case B recombination and cooling rates of Hui & Gnedin (1997) and the photoionization cross-sections of Verner et al. (1996).

We consider two possibilities for the spectra of H I ionizing sources (and hence IGM temperatures following H I reionization) which should bracket the plausible range of possibilities. Following Hui & Haiman (2003), we base the first model on the very massive, metal-free stellar spectrum presented by Bromm, Kudritzki & Loeb (2001): $J_\nu \propto \nu$ for $\nu_{\rm H\,I} \leq \nu \leq \nu_{\rm He\,I}$, $J_\nu \propto \nu^0$ for $\nu_{\rm He\,I} \leq \nu \leq \nu_{\rm He\,II}$ and $J_\nu \propto \nu^{-4.5}$ for $\nu \geq \nu_{\rm He\,II}$, where $\nu_{\rm i}$ corresponds to the ionization threshold for species *i*. This spectrum is very hard, and will result in significant H I *and* He II photoheating during reionization. Note that since this model reionizes He II at high redshift, it is inconsistent with He II reionization completing at lower redshift unless the spectrum softens significantly and He III partially recombines again before $z \simeq 3$. The second model, a more conventional Population II spectrum, is based on the model of Leitherer et al. (1999) for a galaxy of age 500 Myr with a continuous star formation rate, a Salpeter IMF and metallicity $Z = 0.2\,Z_\odot$. In this instance $J_\nu \propto \nu^{-3}$ below the H I ionization threshold. This soft spectrum does not significantly reionize He II. Note that if the IGM is instead reionized by an earlier round of quasar activity, or if quasars are responsible for primarily driving H I reionization (Volonteri & Gnedin 2009, but see Srbinovsky & Wyithe 2007; Willott et al. 2010), the ionizing spectrum and resulting IGM thermal history will be intermediate between these two cases.

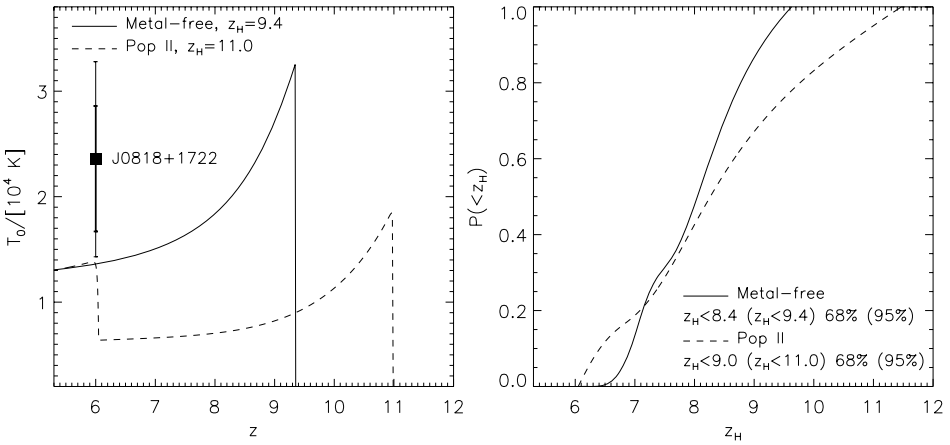

Figure 9. Left-hand panel: the IGM temperature at mean density, T_0, in a localized patch of the IGM as a function of redshift for two different instantaneous reionization models. The solid curve displays the temperature evolution following H I and He II reionization at $z_H = z_{He} = 9.4$ by a very hard spectrum, representative of massive metal-free stars. The dashed curve assumes $z_H = 11.0$ and corresponds to a much softer spectrum representative of Population II sources. Note that He II reionization in this model is caused by the quasar itself at $z_{He} = 6.05$ (i.e. $t_Q = 10^7$ yr before $z = 6$). The filled square shows our constraint on the temperature in the proximity zone of SDSS J0818+1722. The thick (thin) error bars represent the 68 (95) per cent confidence interval around the median. Right-hand panel: the cumulative probability distribution for the redshift of H I reionization around SDSS J0818+1722 we derive from the two different models. Our temperature measurement is consistent with H I reionization occurring at $z_H < 9.0$ ($z_H < 11.0$) at 68 (95) per cent confidence for the Population II model, and $z_H < 8.4$ ($z_H < 9.4$) for the metal-free spectrum. Note, however, that harder ionizing spectra would increase both of these upper limits.

To account for filtering of the ionizing radiation as it propagates through optically thick gas, we assume the mean-free path for ionizing photons in a clumpy IGM has a frequency dependence $\lambda_\nu \propto \nu^{1.5}$ (Zuo & Phinney 1993; Miralda-Escudé 2003). This hardens the intrinsic spectra by $\alpha_s \rightarrow \alpha_s - 1.5$, and is intermediate between the optically thin case and the spectral modification expected for a uniform, optically thick IGM, $\alpha_s \rightarrow \alpha_s - 3$ (Abel & Haehnelt 1999). Lastly, we normalize both models to give an H I photoionization rate of $\Gamma_{HI} = 1.9 \times 10^{-13}\,s^{-1}$, which is consistent with the upper limit inferred from the observed Lyα forest opacity at $z = 6$ (Bolton & Haehnelt 2007c).

Two example thermal histories are displayed in the left-hand panel of Fig. 9, along with our T_0 constraint for the proximity zone of SDSS J0818+1722. The thick (thin) error bars correspond to the 68 (95) per cent confidence intervals around the median T_0. We assume instantaneous H I and He II reionization for the metal-free spectrum at $z_H = z_{He} = 9.4$ (solid curve). The Population II model (dashed curve) instead has $z_H = 11.0$. To account for He II photoheating by the hard spectrum of the quasar itself we also include He II reionization at $z_{He} = 6.05$ (i.e. $t_Q = 10^7$ yr before $z = 6$) for a quasar-like spectrum with $J_\nu \propto \nu^{-1.5}$ (hardened to $J_\nu \propto \nu^0$). This leads to a temperature boost during He II reionization of $\sim$8000 K, consistent with the results of our RT models.

We stress that instantaneous reionization is an unphysical model for the ionization and thermal history of the entire IGM during inhomogeneous H I reionization. However, our temperature constraint is for the volume within 33 Mpc of a bright quasar. If photoheating is the dominant heating process during the epoch of reionization, we should nevertheless obtain an upper limit on the redshift when the bulk of the reionization occurred around the quasar in this approximation. Numerical simulations indicate such biased regions are amongst the first patches of the IGM to be reionized (e.g. Iliev et al. 2006; Shin, Trac & Cen 2008; Finlator et al. 2009). A constraint on z_H within 33 Mpc SDSS J0818+1722 might then also give a constraint for the *global* onset of reionization. However, a joint analysis with other observational constraints including the Lyα forest opacity (Fan et al. 2006) and the electron scattering optical depth (Dunkley et al. 2009) will be required for a consistent interpretation of the temperature data (Choudhury & Ferrara 2006; Pritchard et al. 2009). In this work, *our reionization constraints apply to the IGM around SDSS J0818+1722 only*.

The right-hand panel of Fig. 9 displays the cumulative probability distribution for the redshift of H I reionization, $P(< z_H)$, for these two models using our $z = 6$ temperature measurement. Assuming He II reionization by the quasar boosts the IGM temperature by $\sim$8000 K, we find the Population II model is consistent with H I reionization occurring at $z_H < 9.0$ ($z_H < 11.0$) at 68 (95) per cent confidence. However, a harder quasar spectrum ($\alpha_s < 1.5$), and hence a larger temperature boost during He II reionization, would weaken this constraint considerably. On the other hand, if H I and He II reionization around the quasar are driven primarily by a generation of metal-free stars with hard spectra, we find $z_H < 8.4$ ($z_H < 9.4$) at 68 (95) per cent confidence.

We caution that the astrophysical uncertainties in the modelling presented here are significant; the exact spectrum of the ionizing sources is rather uncertain at $z > 6$. Harder ionizing spectra would increase these upper limits. This study nevertheless demonstrates that even a single temperature measurement at $z = 6$ can give an interesting constraint on the redshift of H I reionization around SDSS J0818+1722. Although computationally expensive, a larger set of simulations will allow for a much more refined calibration of the measurement. Another obvious extension is to analyze additional lines-of-sight (e.g. Becker et al. 2005). Independent measurements along separate lines-of-sight will aid in reducing the statistical uncertainty on any averaged measurement. The amount of scatter in the temperature from one line-of-sight to the next may also provide some insight into the topology and timing of H I reionization globally. Improving existing constraints on the IGM thermal history at $z < 4.5$, as well as adding to constraints at higher redshift, will thus provide extremely valuable insight into the epochs of both He II and H I reionization.

7 CONCLUSIONS

In this paper, we present the first direct measurements of the IGM temperature around a quasar at $z = 6$. We use a combination of high-resolution hydrodynamical simulations combined with a line-of-sight RT implementation to model the thermal state of the IGM in a quasar proximity zone. Previous theoretical studies have suggested that the IGM temperature close to a quasar will be sensitive to the prior ionization state of hydrogen and helium, as well as the intrinsic quasar spectrum (Miralda-Escudé & Rees 1994; Bolton & Haehnelt 2007b; Lidz et al. 2007). We demonstrate here that the Doppler parameter CPDF obtained from the quasar proximity zone provides a sensitive probe of the IGM temperature, enabling us to probe the thermal and ionization history of the IGM within the quasar's vicinity.

Our observational data set consists of a single high-resolution, high S/N Keck/HIRES spectrum of the quasar SDSS J0818+1722 at $z = 6$. We perform identical Voigt profile analyses on the observed and synthetic data to obtain the Doppler parameter CPDF. The simulations, which use a range of self-consistent thermal histories, are used to calibrate the linewidth measurements obtained from the observational data. After verifying our method for obtaining temperature constraints, we proceed to obtain $T_0 = 23\,600 \pm ^{5000}_{6900}$ K ($\pm^{9200}_{9300}$ K) at 68 (95) per cent confidence around the median for the proximity zone of SDSS J0818+1722. Due to the small size of the data set, the error bars are dominated by statistical uncertainty; we have verified that the most important systematic uncertainties remain small in comparison. We find our simulations are also in good agreement with the rapid evolution observed in proximity zone sizes at $z \simeq 6$ when adopting temperatures consistent with our direct constraint. This provides a consistency check of our results, and further suggests that the IGM in the environment of bright quasars is highly ionized by $z \simeq 6$ (Wyithe et al. 2008; Maselli et al. 2009).

Finally, we use our temperature constraint to explore the implications for the IGM reionization history around SDSS J0818+1722 (e.g. Theuns et al. 2002a; Hui & Haiman 2003). We consider two simple models for the thermal history assuming that reionization in the proximity zone occurred instantaneously. Assuming H I reionization around SDSS J0818+1722 is driven by ionizing sources with soft spectra, typical of Population II stars, we infer $z_{\rm H} < 9.0\,(11.0)$ at 68 (95) per cent confidence if photoheating is the dominant heating mechanism during the epoch of reionization. However, if the IGM is instead reionized by a population of massive metal-free stars, characterized by very hard ionizing spectra, we obtain a tighter limit of $z_{\rm H} < 8.4\,(9.4)$. An H I reionization epoch beginning at higher redshifts produces temperatures which are too low with respect to our constraint unless the H I ionizing sources or the quasar itself have spectra significantly harder than typically assumed.

ACKNOWLEDGMENTS

We thank Benedetta Ciardi and Peng Oh for valuable conversations during the course of this work. The hydrodynamical simulations used in this work were performed using the Darwin Supercomputer of the University of Cambridge High Performance Computing Service (http://www.hpc.cam.ac.uk/), provided by Dell Inc. using Strategic Research Infrastructure Funding from the Higher Education Funding Council for England. JSB acknowledges the support of an ARC Australian postdoctoral fellowship (DP0984947), and GDB thanks the Kavli foundation for financial support.

REFERENCES

Abel T., Haehnelt M. G., 1999, ApJ, 520, L13
Alvarez M. A., Abel T., 2007, MNRAS, 380, L30
Becker R. H. et al., 2001, AJ, 122, 2850
Becker G. D., Sargent W. L. W., Rauch M., 2005, in Williams P., Shu C.-G., Menard B., eds, IAU Colloq. 199: Probing Galaxies through Quasar Absorption Lines. Cambridge Univ. Press, Cambridge, p. 357
Becker G. D., Rauch M., Sargent W. L. W., 2007, ApJ, 662, 72
Bolton J. S., Becker G. D., 2009, MNRAS, 398, L26
Bolton J. S., Haehnelt M. G., 2007a, MNRAS, 381, L35
Bolton J. S., Haehnelt M. G., 2007b, MNRAS, 374, 493
Bolton J. S., Haehnelt M. G., 2007c, MNRAS, 382, 325
Bolton J., Meiksin A., White M., 2004, MNRAS, 348, L43
Bolton J. S., Oh S. P., Furlanetto S. R., 2009, MNRAS, 395, 736
Bromm V., Kudritzki R. P., Loeb A., 2001, ApJ, 552, 464
Carilli C. L. et al., 2010, ApJ, in press (arXiv:1003.0016)
Cen R., McDonald P., Trac H., Loeb A., 2009, ApJ, 706, L164
Choudhury T. R., Ferrara A., 2006, MNRAS, 371, L55
Croton D. J., 2009, MNRAS, 394, 1109
Dijkstra M., Lidz A., Hui L., 2004, ApJ, 605, 7
Dunkley J. et al., 2009, ApJS, 180, 306
Eisenstein D. J., Hu W., 1999, ApJ, 511, 5
Fan X., Narayanan V. K., Strauss M. A., White R. L., Becker R. H., Pentericci L., Rix H.-W., 2002, AJ, 123, 1247
Fan X. et al., 2006, AJ, 132, 117
Faucher-Giguère C.-A., Prochaska J. X., Lidz A., Hernquist L., Zaldarriaga M., 2008, ApJ, 681, 831
Finlator K., Özel F., Davé R., Oppenheimer B. D., 2009, MNRAS, 400, 1049
Furlanetto S. R., Oh S. P., 2008, ApJ, 681, 1
Furlanetto S. R., Oh S. P., 2009, ApJ, 701, 94
Gunn J. E., Peterson B. A., 1965, ApJ, 142, 1633
Haardt F., Madau P., 2001, in Neumann D. M., Tran J. T. V., eds, Clusters of Galaxies and the High Redshift Universe Observed in X-rays, preprint (arXiv:astro-ph/0106018)
Haehnelt M. G., Steinmetz M., 1998, MNRAS, 298, L21
Haehnelt M. G., Natarajan P., Rees M. J., 1998, MNRAS, 300, 817
Hui L., Gnedin N. Y., 1997, MNRAS, 292, 27
Hui L., Haiman Z., 2003, ApJ, 596, 9
Hui L., Rutledge R. E., 1999, ApJ, 517, 541
Iliev I. T., Mellema G., Pen U.-L., Merz H., Shapiro P. R., Alvarez M. A., 2006, MNRAS, 369, 1625
Kelson D. D., 2003, PASP, 115, 688
Kim T.-S., Bolton J. S., Viel M., Haehnelt M. G., Carswell R. F., 2007, MNRAS, 382, 1657
Kirkman D. et al., 2005, MNRAS, 360, 1373
Komatsu E. et al., 2009, ApJS, 180, 330
Leitherer C. et al., 1999, ApJS, 123, 3
Li Y. et al., 2007, ApJ, 665, 187
Lidz A., McQuinn M., Zaldarriaga M., Hernquist L., Dutta S., 2007, ApJ, 670, 39
Lidz A., Faucher-Giguere C., Dall'Aglio A., McQuinn M., Fechner C., Zaldarriaga M., Hernquist L., Dutta S., 2009, ApJ, submitted (arXiv:0909.5210)
McDonald P., Miralda-Escudé J., Rauch M., Sargent W. L. W., Barlow T. A., Cen R., Ostriker J. P., 2000, ApJ, 543, 1
McDonald P., Miralda-Escudé J., Rauch M., Sargent W. L. W., Barlow T. A., Cen R., 2001, ApJ, 562, 52
McQuinn M., Lidz A., Zaldarriaga M., Hernquist L., Hopkins P. F., Dutta S., Faucher-Giguère C.-A., 2009, ApJ, 694, 842
Madau P., Meiksin A., 1994, ApJ, 433, L53
Martini P., 2004, in Ho L. C., ed., Coevolution of Black Holes and Galaxies. Cambridge Univ. Press, Cambridge, p.169
Maselli A., Gallerani S., Ferrara A., Choudhury T. R., 2007, MNRAS, 376, L34
Maselli A., Ferrara A., Gallerani S., 2009, MNRAS, 395, 1925
Meiksin A., Tittley E. R., Brown C. K., 2010, MNRAS, 401, 77

Mesinger A., Haiman Z., 2007, ApJ, 660, 923
Miralda-Escudé J., 2003, ApJ, 597, 66
Miralda-Escudé J., Rees M. J., 1994, MNRAS, 266, 343
Olive K. A., Skillman E. D., 2004, ApJ, 617, 29
Pawlik A. H., Schaye J., van Scherpenzeel E., 2009, MNRAS, 394, 1812
Peeples M. S., Weinberg D. H., Davé R., Fardal M. A., Katz N., 2009, MNRAS, submitted (arXiv:0910.0256)
Press W. H., Teukolsky S. A., Vetterling W. T., Flannery B. P., 1992, in Press W. H., Teukolsky S. A., Vetterling W. T., Flannery B. P., eds, Numerical Recipes in FORTRAN. The Art of Scientific Computing. Cambridge Univ. Press, Cambridge
Pritchard J. R., Loeb A., Wyithe J. S. B., 2009, MNRAS, submitted (arXiv:0908.3891)
Rauch M. et al., 1997, ApJ, 489, 7
Reichardt C. L. et al., 2009, ApJ, 694, 1200
Ricotti M., Gnedin N. Y., Shull J. M., 2000, ApJ, 534, 41
Schaye J., Theuns T., Leonard A., Efstathiou G., 1999, MNRAS, 310, 57
Schaye J., Theuns T., Rauch M., Efstathiou G., Sargent W. L. W., 2000, MNRAS, 318, 817
Shin M., Trac H., Cen R., 2008, ApJ, 681, 756
Songaila A., 2004, AJ, 127, 2598
Springel V., 2005, MNRAS, 364, 1105
Springel V., Hernquist L., 2003, MNRAS, 339, 289
Srbinovsky J. A., Wyithe J. S. B., 2007, MNRAS, 374, 627
Telfer R. C., Zheng W., Kriss G. A., Davidsen A. F., 2002, ApJ, 565, 773
Theuns T., Leonard A., Efstathiou G., Pearce F. R., Thomas P. A., 1998, MNRAS, 301, 478
Theuns T., Schaye J., Haehnelt M. G., 2000, MNRAS, 315, 600
Theuns T., Schaye J., Zaroubi S., Kim T., Tzanavaris P., Carswell B., 2002a, ApJ, 567, L103
Theuns T., Zaroubi S., Kim T.-S., Tzanavaris P., Carswell R. F., 2002b, MNRAS, 332, 367
Trac H., Cen R., Loeb A., 2008, ApJ, 689, L81
Tytler D. et al., 2004, ApJ, 617, 1
Valageas P., Schaeffer R., Silk J., 2002, A&A, 388, 741
Verner D. A., Ferland G. J., Korista K. T., Yakovlev D. G., 1996, ApJ, 465, 487
Volonteri M., Gnedin N. Y., 2009, ApJ, 703, 2113
Volonteri M., Rees M. J., 2006, ApJ, 650, 669
Walter F., Carilli C., Bertoldi F., Menten K., Cox P., Lo K. Y., Fan X., Strauss M. A., 2004, ApJ, 615, L17
White R. L., Becker R. H., Fan X., Strauss M. A., 2003, AJ, 126, 1
Willott C. J. et al., 2010, AJ, 139, 906
Wyithe J. S. B., Bolton J. S., Haehnelt M. G., 2008, MNRAS, 383, 691
Zaldarriaga M., 2002, ApJ, 564, 153
Zuo L., Phinney E. S., 1993, ApJ, 418, 28

This paper has been typeset from a TEX/LATEX file prepared by the author.